JAPANESE [JP,08-126003,A]

CLAIMS DETAILED DESCRIPTION PRIOR ART EFFECT OF THE INVENTION TECHNICAL PROBLEM MEANS OPERATION EXAMPLE DESCRIPTION OF DRAWINGS DRAWINGS

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]
[Claim 1] (a) While detecting the binary image in which the edge location which is the boundary of reception and this image data to a pixel value change about the image data of two-dimensional size is shown The 1st [ which corresponds the 1st edge positional information which shows the direction component of said edge location from this binary image ] detection means which detects for every pixel, (b) The 2nd detection means which detects the 2nd edge positional information which reception and a direction component carry out grouping of said 1st edge positional information relevant to mutual for said 1st edge positional information, and shows the edge location by which grouping was carried out, (c) A means to ask for the transition location of a pixel value by referring to each pixel value in said image data area corresponding to reception and said 2nd edge positional information for said 2nd edge positional information, (d) Image processing system which has a coding means to encode the difference of the edge location specified to said 2nd edge positional information and said 2nd edge positional information, and said transition location.
[Claim 2] The image processing system characterized by including the geometric information which shows the die length of the edge by which grouping was carried out to said 2nd edge positional information, a direction, and a starting position in the 1st term of a claim.
[Claim 3] A means to ask for the transition location of said pixel value in the 1st term of a claim is an image processing system characterized by being carried out by comparing the average of each pixel value and each pixel value in said image data area.
[Claim 4] (a) A detection means to detect the edge positional information which shows the edge location which is the boundary of reception and this image data to a pixel value about the image data of two-dimensional size, (b) A coding means to encode said edge positional information, and an image reconstruction means to perform image reconstruction based on the edge positional information by which (c) coding was carried out, (d) — the difference of said image data and the playback image data reproduced by said playback means — a means to encode a value, the image based on (e) edge positional information, and difference — the image processing system [claim 5] which has the means which raises image quality gradually by carrying out sequential transmission of the value nonlinear sampling said whose coding means used edge positional information in the 4th term of a claim — said difference — the image processing system characterized by including a means to encode a value.
[Claim 6] (a) Detect the binary image in which the edge location which is the boundary of reception and this image data to a pixel value change about the image data of two-dimensional size is shown. (b) The 1st edge positional information which shows the direction component of said edge location from said binary image is detected for every corresponding pixel. (c) Reception and a direction component carry out grouping of said 1st edge positional information relevant to mutual for said 1st edge positional information. The 2nd edge positional information which shows the edge location by which grouping was carried out is detected. (d) It asks for the transition location of a pixel value by referring to each pixel value in said image data area corresponding to reception and said 2nd edge positional information for said 2nd edge positional information. Edge positional information of the (e) above 2nd, The image-processing approach of

having the step which encodes the difference of the edge location specified to said 2nd edge positional information, and said transition location.

---

[Translation done.]

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Description of the Prior Art] In order that International Organization for Standardization (ISO) might answer the need for the general approach of encoding an animation by the low bit rate and might build the video coding standard about various applications, such as a video telephone and mobile communications, it formed MPEG-4 (Moving Pictures Expert Group, Phase 4) in 1993. The main purposes of many picture compression encoding methods tend to perform high playback of the fidelity about a subject-copy image with high-pressure shrinking percentage as much as possible. The criteria of fidelity of being used for the design of the compression encoding method play a big role on the engine performance. The criteria of fidelity of generally being used are an average square error (MSE). The main special features of MSE are that the mathematical count is easy and that the small value of MSE actually corresponds to the playback image of high quality sensuously. Since a final judgment of the reproduced image is made by human being's eyes, the latter is important.
[0002] The various image coding techniques of having MSE as criteria of fidelity, such as conversion coding, have been developed. Although these techniques bring about a comparatively quality playback image by the bit rate beyond a pixel, 1.0 bits /, and it, they are one of these and produce often more special visible degradation, for example, block distortion, edge dotage, etc. in a low bit rate. The new class of the image encoding method well learned as a second generation coding technique has been developed over the past ten years or more. These approaches tend to prove upgrading of the technique of MSE orientation [ bit rate / low-speed ] very much, and tend to draw an image by expressing a true stereo like a border line or texture in a compact further. So, the second generation encoding method is expected to attain high-pressure shrinking percentage on the criteria of the fidelity to which human being's eye which is not yet realized pointed.
[0003] Image coding based on sketch drawing is considered as a typical means of this category, and has been quoted as reference. This approach characterizes according to a profile line-extraction process, and results in bringing about a harmful distortion in sketch drawing playback to a local noise, by detection of a pixel unit, since it is brittle.
[0004]
[Problem(s) to be Solved by the Invention] Moreover, a subject-copy image is conventionally disassembled into two components in addition to above-mentioned technique, and the technique of encoding those components is known. This technique disassembles a subject-copy image into a primary component (primary component) including edge information, and the loose smooth component (secondary component) of brightness change. The idea of decomposition coding was produced from 2 component model by John (Yan) and SAKURISON (Sakrison), and some practical approaches have been studied since then.
[0005] All the approaches by these decomposition coding have the impact for the image quality of a playback image, as a result the whole engine performance with this big focusing on an edge extract, i.e., detection, an expression, and coding. Chain coding is used for a great portion of technique, in order to perform edge detection of a pixel unit like a peak point trace, i.e., an edge

trace, and to encode the sequence of an edge location. However, it is pointed out that the problem of a closed curve and dealing with a local noise are difficult, and, as for the above-mentioned technique, cause the result of interruption of an edge, a location error, and inaccurate reinforcement. Moreover, since it was thought that the information on a context exists only in the field restricted very much, there was no method of distinguishing the border line corresponding to an objective boundary and the border line which is not so.

[0006] The purpose of this invention solves the technical problem of the above-mentioned conventional technique, and offers the image processing system which can acquire edge positional information from image data, and its approach by hierarchical edge detection. Other purposes of this invention offer the image of high quality while they introduce each step of hierarchical edge detection, i.e., unit edge detection, macro edge detection, and local adjustment MENTO and express edge information in a compact.

[0007] Other purposes of this invention offer the image processing system which fidelity is excellent in in coding of image data, and encodes the image data based on high-pressure shrinking percentage, and its approach. Other purposes of this invention decompose image data into primary image components and secondary image components, and offer the image processing system which can reproduce the primary image by which edge orientation was carried out from primary image components, and its approach.

[0008]

[Means for Solving the Problem] It has the following means, in order to solve the above-mentioned technical problem. While the image processing system concerning invention of the 1st of this application detects the binary image in which the edge location which is the boundary of reception and this image data to a pixel value change about the image data of two-dimensional size is shown The 1st [ which corresponds the 1st edge positional information which shows the direction component of said edge location from this binary image ] detection means which detects for every pixel, The 2nd detection means which detects the 2nd edge positional information which reception and a direction component carry out grouping of said 1st edge positional information relevant to mutual for said 1st edge positional information, and shows the edge location by which grouping was carried out, A means to ask for the transition location of a pixel value by referring to each pixel value in said image data area corresponding to reception and said 2nd edge positional information for said 2nd edge positional information. It has a coding means to encode the difference of the edge location specified to said 2nd edge positional information and said 2nd edge positional information, and said transition location.

[0009] Moreover, a detection means to detect the edge positional information the image processing system concerning the 2nd invention indicates the edge location which is the boundary of reception and this image data to a pixel value about the image data of two-dimensional size to be, A coding means to encode said edge positional information, and an image reconstruction means to perform image reconstruction based on the encoded edge positional information, the difference of said image data and the playback image data reproduced by said playback means — a means to encode a value, the image based on edge positional information, and difference — it has the means and ** which raise image quality gradually by carrying out sequential transmission of the value. Since the application by this invention is a thing relevant to MPEG-4 area, the video telephone which deals with several sorts of bodies, and its pocket communication link are desirable as target application. The test sequence of a bust is used for simulation by such intention.

[0010]

[Function] According to this invention, by having used hierarchical edge detection (the unit edge detection shown in an example, macro edge detection, and step of local adjustment MENTO), the coding technique of edge contrast directivity can be offered and coding processing of the outstanding image data can be performed in consideration of the moral vision property of human being's vision system.

[0011]

[Example] Hereafter, the example of the image processing system of this invention is explained to a detail with reference to a drawing. Drawing 1 is the block diagram showing the configuration

of decomposition coding concerning the image processing system of this example. As shown in this drawing, a subject-copy image is supplied to the edge extract section 10, and the edge information about the border line of the body in an image etc. is extracted here. The extracted edge information is supplied to the data optimization section 12, and in order to attain high-pressure shrinkage, edge information-redundancy nature is removed. In this way, the primary component (primary component) 14 about the edge information extracted from the subject-copy image is obtained.

[0012] Moreover, the primary component 14 is supplied to the image reconstruction section 16, and the primary image 18 is reproduced based on a primary component. and the reproduced primary image 18 — difference — a subject-copy image and difference ask in a vessel 20 — having — this difference — a value is supplied to the differential-encoding section 22. here — difference — DCT processing by the adjustable size block is performed about a value, and the encoded smooth component (secondary component) 24 is obtained. Drawing 2 is drawing showing the hierarchization edge extract process in the edge extract section of drawing 1 . The process of the edge extract from a subject-copy image is performed through each of the Laplacian filter 100, the unit edge detecting element 102, the macro edge detecting element 104, the local adjustment MENTO section 106, and the on-the-strength count section 108.

[0013] First, as a subject-copy image, the image data (brightness component) of 480*704-pixel size is supplied to the Laplacian filter 100, and well-known Laplacian processing is performed. That is, secondary differential data in which the change about each pixel is shown are called for. In addition, the image data of a color difference component is the size sampled by one half in 480*704, and these data are used in the on-the-strength count section 108 mentioned later. Next, the data by which Laplacian processing was carried out are supplied to the unit edge detecting element 102, and the binary image in which an exact edge location is shown is obtained by using mu+K-sigma for a threshold. Here, mu, sigma, and K are an average, the standard deviation of derivative space, and a multiplier, respectively. The concept of an edge location is used as a thing showing the boundary of the change, when the brightness of each pixel in image data changes steeply and continuously.

[0014] And it matches about the pixel which shows the edge location in a subject-copy image using the segment pattern in which eight directions as shown in drawing 3 are shown. The pattern for matching is shown by Template Tn (n= 0, 1, .. 7), and each entry in (j, k) is expressed by tn (j, k). the sub field which consists lambda (x y) of 5*5pixellambda (x+j, y+k) — it is — making (j, k= 0, 1, 2, 3, 4) . The cross-correlation CRn (x y) between Tn and lambda (x y) is calculated by the degree type.

[Equation 1]

$$CRn\ (x,\ y) = \sum_{j=0}^{4} \sum_{k=0}^{4} tn\ (j,\ k) \times \lambda\ (x+j,\ y+k) \cdots (1)$$

[0015] Therefore, CRn (x y) is equal to 7, or if n which is seven or more exists, a flag will stand on the coordinate (x y) of n bit plane. This shows that it was chosen as a matching pattern which Template Tn is a coordinate (x y) and calls a unit edge. For example, in matching about the pixel which shows a certain edge location, the pixel concerned used as a processing object is located at the core (location of three-line three trains) of a template, and if the neighboring pixel of the pixel concerned is located horizontally, template T four will be chosen, and the edge location which is the pixel concerned is processed as a thing with the direction component of T four. In addition, among drawing, the numeric value of "1" and "2" shows weighting of matching, and "2" is processing it as a thing heavier than "1" by this example.

[0016] Next, grouping, i.e., macro-izing, is performed by the macro edge detecting element 104 about the pixel by which the unit edge was detected. As mentioned above, although a unit edge is prescribed by the template of eight directions, it is connected to the macro edge on which each of these unit edges are specified in the 16 directions in a continuous phase.

[0017] If the pixel concerned which should be processed shall be located in one line, five trains, and (1, 5) supposing the pixel matrix of five line *9 train, the direction of 16 specified on a macro edge the location of (5, 5) of right under [ the ] — direction "0" — carrying out — and order —

(5, 4), and ... (5 3) (5 1) — respectively — a direction — "1" and "2 "and" 4" — carrying out — moreover, order — (4, 1), (3, 1), and (2, 1) — respectively — a direction — it is referred to as "5 "and" 7." the same — the order from direction "0" — right-hand side — (5, 6), and ... (5 7) (5 9) — respectively — a direction — "15" and "14 "and" 12" — carrying out — moreover, order — (4, 9), (3, 9), (2, 9), and (1, 9) — respectively — a direction — it is referred to as "11" and "10 "and" 8." The direction n of a unit edge corresponds to N specified at a ceremony (2), and N is the core of the direction of a base that search processing for connection is performed.

[Equation 2] N= 2n ... (2)

[0018] The direction of the connection in a macro edge must be detected out of the direction of three candidates, i.e., N, N−1, and N+1. For example, if a certain unit edge is a template T1, the unit edge which is N= 2, therefore should be connected from the direction of 1, 2, and 3 will be detected. A unit edge must be connected in the direction which can extract a macro edge with the longest criteria of selection. In the direction of each candidate, it is determined in each node segmented by the unit length Lunit (4 pixels is said) whether a macro edge is connected. If the flag of the bit side n, n−1, and n+1 arises near a node, a macro edge will be extended till a node. It is dependent on what is carried out what magnitude of a field counts for such decision. In addition, the unit edge of the pixel which will be located on the macro edge is eliminated after connection processing termination of a macro edge.

[0019] Here, with reference to drawing 4 and drawing 5 , the concrete example of macro edge detection is explained. A macro edge is detected by connecting a unit edge in the following procedures. First, the unit edge developed on n−bit plane is considered to be seemingly developed by N−bit plane so that it may correspond to (2) types. Supposing the flag stands on N−bit plane now, it will consider as the starting point lambda of the macro edge which detects the location after this. It is node piN and L about the point searched in case L unit edges are connected in the direction N from the starting point lambda. A definition is given and it is lambda in Lunit=4, piN, and L. Physical relationship is shown in drawing 4 . At this example, it is piN and L about a search aperture. It will consider as the 3x3−pixel field made into a core, and if the flag stands on this field on N, N−1, and an N+1−bit plane, sequential increment of the number L of connection is carried out, a search will be repeated and the node of the last whose connection was completed will be made into the terminal point of the candidate macro edge of Direction N. The candidate macro edge of a direction N−1 and a direction N+1 is also detected from the same starting point (if the flag stands into the search aperture on N−2 or N−bit plane by the search of a direction N−1 in these cases, it will be judged as connection), and, finally the number L of connection detects the greatest thing as a macro edge in these 3 candidate. Supposing two or more candidates with the maximum number L of connection exist, this number will choose the larger one, using the total of the flag in a search aperture as secondary scale. Next, the example of connection of a unit edge is shown in drawing 5 . The minimum field divided by the ruled line is a pixel, and each block diagram expresses the same subregion on an image. Let the first location scans N bit plane and the flag stands be the starting point lambda of a macro edge. In this example, lambda is called for on 2−bit plane. Therefore, the search for connection will search a direction 3 by the search of a direction 1, and 2−bit plane by 0−bit plane at the search of directions 1, 2, and 3, and 4−bit plane. piN in case the number L of connection becomes max in the search of directions 1, 2, and 3, respectively, and L A location is shown in the lower berth of drawing 5 . Finally at this example, it is pi 1 and 3. It becomes the terminal point of the macro edge to detect.

[0020] In this way, detection of the macro edge about each unit edge is performed, and the starting position of the macro edge by which grouping was carried out, the direction of either of 16, and die length are obtained. Next, in the local adjustment MENTO section 106, the rectangle field which wraps a macro edge in predetermined die−length Lext (this example 7 pixels) is specified as an edge belt E using the called−for macro edge. Drawing 6 is the example of a belt edge and sets a shaft perpendicular to a shaft parallel to a macro edge to p and q, respectively. And the pixel value on the edge belt E is expressed as epsilon (x y).

[0021] Generally, an actual edge can be assumed to be what exists along with the macro edge of the edge belt E. In order to locate an actual edge correctly, change (8−bit gradation) of the gray

level of a pixel is inspected in accordance with the shaft q perpendicular to a macro edge. The step is explained.

(i) The average phi of the gray level of all the pixels in an edge belt is calculated first.

(ii) If it is smaller than the average psi, 0 is marked on the pixel corresponding to epsilon (x y), and if the gray level of each pixel is larger than the average psi, it will mark 1 about each pixel of the edge belt E.

(iii) An actual edge is located to the place which the transition to 0 to 1, or 1-0 produces (default"0" which expresses the pixel which does not produce such transition about the shaft on a macro edge is used).

(iv) By calculating respectively the average value about the pixel marked by the both sides of 0 and 1, the edge profile which has an ideal step function is approximated, and it is the reinforcement delta 0 of the lower one. Reinforcement delta 1 of the higher one It obtains, respectively.

[Equation 3]

$$\delta_0 = \frac{1}{\tau_0} \sum_{(p,q)} \sum_{<\phi} \varepsilon(p, q) \qquad \cdots(3)$$

$$\delta_1 = \frac{1}{\tau_1} \sum_{(p,q)} \sum_{\geq\phi} \varepsilon(p, q) \qquad \cdots(4)$$

[0022] Here, it is tau 0. tau 1 The number of pixels with which are satisfied of the monograph affair of a sum total type (3) and (4) is shown. Drawing 7 shows the example of tic [ SUKEMA ] acquired by the local adjustment MENTO section 106. An actual edge location is pursued by the thick wire. Moreover, the sequence of the pixel of a shaft q= 0 corresponds to a macro edge. In addition, the on-the-strength count section 108 performs count of the above-mentioned sum total type (3) and (4) based on the result of the local adjustment MENTO section 106. Moreover, the on-the-strength count section 108 performs count on the strength with the same said of a color difference component.

[0023] The edge data obtained according to the above process are outputted to the data optimization section 12 from the edge extract section 10 shown in drawing 1 . In order to attain high-pressure shrinking percentage, the data optimization section 12 is removed from the edge data from which redundancy and the information which is not not much important were extracted, and encodes edge data.

(i) Since the actual edge obtained by the local adjustment MENTOROKARUAJASUTOMENTO section 106 has the low pass property which met in the direction of a macro edge as shown in above-mentioned drawing 7 , by carrying out subsampling in the predetermined period Lsub, it reduces data and performs differential encoding about the reduced data.

(ii) -- reinforcement and Weber FEFUNA -- with reference to a brightness difference threshold, this removes the edge which is not not much important in human being's vision sensibility using law. Let deltaI be the brightness difference threshold of an illuminance I. deltaI is prescribed that brightness becomes remarkable, when a brightness difference reaches deltaI or it is exceeded.

[Equation 4]

$$\frac{\Delta I}{I} = \theta \qquad \cdots(5)$$

[0024] The above-mentioned formula can be expressed with a multiplier zeta, being able to assume that it is what deltaI/I gives the magnitude of the vision sensibility E well.

[Equation 5]

$$\Delta E = \zeta \frac{\Delta I}{I} \qquad \cdots(6)$$

It integrates with this and is [Equation 6]. E=zeta'logI ... (7)

In order to apply the Weber FEFUNA method to the encoding method, a formula (5) is

transformed using an original definition. That is, the macro edge with which are satisfied of a bottom type is removed from the data of a primary component.
[Equation 7] delta1−delta0 <=thetapsi ... (8)
[0025] In an actual case, thetay and thetac can be set about brightness and a chrominance, respectively.
[Table 1]

表 1

| カテゴリ | 符号化されるメッセージ | 範囲 | ビット数 |
|---|---|---|---|
| 幾何学的情報 | 始点 | ピクチャサイズ | $\mathrm{Log}_2$（横サイズ）+$\mathrm{Log}_2$（縦サイズ） |
| | 方向 | [0,15] | 4 |
| | Lunit を1単位とした長さ | [1,32]* | 可変長コード |
| ローカルアジャストメント | 実際のエッジ位置と異なる | [−7,7]* | 可変長コード |
| 強度 | クロマ重要度フラッグ | 0又は1 | 1（クロミナンスのみ） |
| | ステップ関数のタイプ | 0又は1 | 1 |
| | 低い方の強度：$\delta_0$ | [0,255] | 6 |
| | コントラスト：$\delta_1-\delta_0$ | [0,255] | 可変長コード |
| ノート：* 画像サイズ及び/又は手段による | | | |

[0026] Table 1 shows the message for encoding the macro edge about a primary component. As geometric information on a macro edge, the starting point of a macro edge, a direction, and the die length of a unit have the predetermined range and the number of bits, and are encoded. Moreover, about local adjustment MENTO, the difference of a macro edge and an actual edge location is encoded. Moreover, in this example, the element of the both sides of brightness and a chrominance is encoded. Geometric information and geometric local adjustment MENTO are obtained using a luminance element, and reinforcement other than another side and the flag showing the semantics of a chroma (color) is calculated for every color element.
[0027] Next, the image reconstruction section 16 is explained. The primary component 14 contains only the reinforcement of edge associated data, i.e., geometric information, and each macro edge. So, a certain kind of interpolation/extrapolation must be used in order to predict the gray level in fields other than an edge belt. Reconstructive processing draws the pixel in each edge belt using (i) geometry information, local adjustment MENTO, and reinforcement (namely, contrast). About local adjustment MENTO, interpolation of the actual edge location between adjoining sampling points is carried out in linearity. In this way, each pixel in an edge belt can give a gray level depending on the side in which an actual edge is located.
(ii) The reference pixel of eight directions is used for interpolation, omegai which predicts the gray level on pixel criteria, and alphai are made into distance from the pixel predicted to be a reference pixel, and phi on Direction i shows each (refer to drawing 8 ). Other reference pixels do not exist between phi and omega i. omegai is made into the gray level in reference pixel omegai, and the gray level of the pixel which is shown by phi and which is predicted is called for from a bottom type.
[Equation 8]

$$\bar{\phi} = \frac{1}{\sum_{i=0}^{7} \alpha i^{-1}} \sum_{i=0}^{7} \bar{\omega} i \quad \alpha i^{-1} \qquad \cdots (9)$$

[0028] Such processing is performed until it results far away from a near edge belt, in order to obtain the smooth change by the gray level. The segmentation in an adjustable size block is performed to the whole image, and interpolation is performed from a small block to a big block.
[0029] Next, the differential-encoding section is explained. Although the primary image 18 reproduced by the image reconstruction section 16 is lacked in the detail of a smooth field, i.e., the loose field of a change on the strength, it offers the depiction which was excellent in consciousness. In order to fill up the detail in this field, the so-called adjustable block-size coding processing on the basis of the edge orientation sensibility of human being's vision system is used. That is, coding by the small block size is performed in the neighborhood of an edge, and coding by another side and the big block size is used when the distance from an edge increases. The combination of a block size 8*8, 16*16, and 32*32 is applied in order to attain high-pressure shrinking percentage. In this way, the quality image defined as the second image with the SUKERA kinky thread tee of SNR can be obtained.
[0030] This technique is coding characterized by nonlinear sampling using edge information, and shows the outline of the processing to drawing 9 . it mentioned above — as ( drawing 1 ) — the difference of the subject-copy image 200 and primary component images 18 — a value — difference — it asks with a vessel 20 — having — this difference — a value is supplied to the nonlinear sampling section 202. Nonlinear sampling to which a block size is changed according to the local property of an image is called the adaptation block encoding method, and the various implementation technique is proposed. Those most are transmitting additionally the information which shows a block size, and the information which shows division of a block. In order that this technique may utilize the information on the edge (after local adjustment MENTO application) which is the coded data of a primary component and may realize nonlinear sampling, it does not need additional information. In this example, a square block (the block of three kinds of magnitude, i.e., 32-pixel S*32 base, 16 pixels * 16 pixels, and 8 pixels * 8 pixels) is used. First, an edge is developed on the bit plane of the same magnitude as an image, and a flag is set in the location where the element of an edge exists. Next, a 32-pixel block [ -32 pixel ] performs linearity sampling. And if the flag stands in the block concerned about each block, it divides into four 16 pixel x 16-pixel blocks and the flag does not stand, suppose that it remains as it is. Similarly, in the next phase, if the flag stands in the block concerned about the 16 pixel x 16-pixel block, it divides into four 8 pixel x 8-pixel blocks and the flag does not stand, suppose that it remains as it is. Thus, improvement in vision evaluation is expectable by changing a block size depending on the distance from an edge.
[0031] Processing after nonlinear sampling is performed one by one by the discrete cosine transform section 204 and multiplier quantization which are generally performed and the capable block judging section 206, the multiplier scan section 208, and the run level coding section 210. It is fundamentally [ as MPEG-1 (Moving Pictures Expert Group Phase-1:ISO/IEC -11172) which is JPEG (Joint Photographic Expert Group;ISO -10918) and the dynamic-image coding standard which are coding using a discrete cosine transform (DCT), for example, a color static-image coding standard, and MPEG-2 (Moving Pictures Expert Group Phase-2:ISO/IEC -13818) ] the same. By the proposal technique, using three kinds (namely, 32 pixel x32 pixel, 16 pixel x16 pixel, and 8 pixels x 8 pixels) of discrete cosine transforms with nonlinear sampling is mentioned to using as difference the discrete cosine transform these criteria of whose are 8 pixel x8 pixels.
[0032] Next, the result of the simulation of this example is shown below. the engine performance of above-mentioned this example — simulating — the intra of MPEG-2 — it compared with the engine performance of image (inside of frame) coding. Simulation conditions are K= 1.0, Lunit=4 pixel, Lext=7 pixel, thetay=0.10, thetac=0.02, and Lsub=4 pixel.
[0033] As shown in drawing 10 , test sequence "Susie" whose pixel size is 704 pixels * 480 lines is used. A color format is 4:2:0 and is 8 bits/pixel. The result of macro edge detection and local

adjustment MENTO is shown in drawing 11 and drawing 12 , respectively. a playback image, i.e., a primary image, and the second image — respectively — drawing 13 and 15 — being shown — them and HPEG- SNR to which an image corresponds 2 intra is shown in Table 2.
[Table 2]

表 2

| 符号化法 | ビットレート ビット/フレーム | SNRS(dB) | | |
|---|---|---|---|---|
| | | Y | Cb | Cr |
| 1次画像 | 15,049 | 22.81 | 36.31 | 34.29 |
| 2次画像 | 57.168(total) | 32.70 | 41.38 | 40.93 |
| MPEG-2 (I-画像) | 81,224 | 31.61 | 44.11 | 43.45 |

The segmentation in the adjustable size block used for the both sides of interpolation processing and differential encoding is shown in drawing 14 . In this example, 255 macro edges extracted from the subject-copy image exist. In addition, as for this drawing (a), the 2nd-step division image and this drawing (c) show the last division image, as for an initial division image and this drawing (b).
[0034] The hierarchical edge detection by this example gives the compact expression of edge information, and, so, a primary image gives the rough outline of a body as shown in drawing 13 , or a scene. Although the compressibility about a primary component was 250:1 or more, image quality was not suitable in itself. Addition of a smooth component attains 70:1 or more compressibility, and raises image quality considerably. From the above-mentioned simulation result, the picture compression encoding method by this example is a low bit rate, is SNR which is equal to image coding in MPEG-two frames (intra), and offers high definition more.
[0035] This invention has the description based on the encoding method using the hierarchical edge detection equipped with differential encoding. This approach is accomplished corresponding to the need about the latest low bit rate / high-pressure shrinking percentage image encoding method. The model used in the example of this invention disassembles an image into the primary component containing an edge element, and the smooth component in which a change on the strength carried out slowly is shown fundamentally. The effectiveness of this invention which let pass with simulation and was obtained is as follows.
[0036] MPEG- which mentioned [ 1st ] the technique of this example above — the engine performance of image coding (I-picture) is improved 2 intra, and MSE based on the precision of another side and a body can match. What mainly contributes to the 2nd at such engine-performance amelioration is the effective sensuous tuning [ express and ] on the basis of the direction and differential encoding, i.e., adjustable block-size coding, of edge information which used hierarchical edge detection. Furthermore, this example enables the gradual transfer to a secondary image from a primary image, it is desirable to application like browsing in an image database etc., and most of another side and second generation coding techniques do not give this description.
[0037]
[Effect of the Invention] According to the image processing system concerning this invention, coding processing of the image data which could offer the method of encoding edge contrast directivity, and was excellent in consideration of the moral vision property of human being's vision system can be performed by having used hierarchy edge detection (the unit edge detection shown in an example, macro edge detection, and step of local adjustment MENTO).

---

[Translation done.]

## PRIOR ART

[Description of the Prior Art] In order that International Organization for Standardization (ISO) might answer the need for the general approach of encoding an animation by the low bit rate and might build the video coding standard about various applications, such as a video telephone and mobile communications, it formed MPEG-4 (Moving Pictures Expert Group, Phase 4) in 1993. The main purposes of many picture compression encoding methods tend to perform high playback of the fidelity about a subject-copy image with high-pressure shrinking percentage as much as possible. The criteria of fidelity of being used for the design of the compression encoding method play a big role on the engine performance. The criteria of fidelity of generally being used are an average square error (MSE). The main special features of MSE are that the mathematical count is easy and that the small value of MSE actually corresponds to the playback image of high quality sensuously. Since a final judgment of the reproduced image is made by human being's eyes, the latter is important.
[0002] The various image coding techniques of having MSE as criteria of fidelity, such as conversion coding, have been developed. Although these techniques bring about a comparatively quality playback image by the bit rate beyond a pixel, 1.0 bits /, and it, they are one of these and produce often more special visible degradation, for example, block distortion, edge dotage, etc. in a low bit rate. The new class of the image encoding method well learned as a second generation coding technique has been developed over the past ten years or more. These approaches tend to prove upgrading of the technique of MSE orientation [ bit rate / low-speed ] very much, and tend to draw an image by expressing a true stereo like a border line or texture in a compact further. So, the second generation encoding method is expected to attain high-pressure shrinking percentage on the criteria of the fidelity to which human being's eye which is not yet realized pointed.
[0003] Image coding based on sketch drawing is considered as a typical means of this category, and has been quoted as reference. This approach characterizes according to a profile line-extraction process, and results in bringing about a harmful distortion in sketch drawing playback to a local noise, by detection of a pixel unit, since it is brittle.

[Translation done.]

## EFFECT OF THE INVENTION

[Effect of the Invention] According to the image processing system concerning this invention, coding processing of the image data which could offer the method of encoding edge contrast directivity, and was excellent in consideration of the moral vision property of human being's vision system can be performed by having used hierarchy edge detection (the unit edge detection shown in an example, macro edge detection, and step of local adjustment MENTO).

[Translation done.]

## TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] Moreover, a subject-copy image is conventionally disassembled into two components in addition to above-mentioned technique, and the technique of encoding those components is known. This technique disassembles a subject-copy image into a primary component (primary component) including edge information, and the loose smooth component (secondary component) of brightness change. The idea of decomposition coding was produced from 2 component model by John (Yan) and SAKURISON (Sakrison), and some practical approaches have been studied since then.
[0005] All the approaches by these decomposition coding have the impact for the image quality of a playback image, as a result the whole engine performance with this big focusing on an edge extract, i.e., detection, an expression, and coding. Chain coding is used for a great portion of technique, in order to perform edge detection of a pixel unit like a peak point trace, i.e., an edge trace, and to encode the sequence of an edge location. However, it is pointed out that the problem of a closed curve and dealing with a local noise are difficult, and, as for the above-mentioned technique, cause the result of interruption of an edge, a location error, and inaccurate reinforcement. Moreover, since it was thought that the information on a context exists only in the field restricted very much, there was no method of distinguishing the border line corresponding to an objective boundary and the border line which is not so.
[0006] The purpose of this invention solves the technical problem of the above-mentioned conventional technique, and offers the image processing system which can acquire edge positional information from image data, and its approach by hierarchical edge detection. Other purposes of this invention offer the image of high quality while they introduce each step of hierarchical edge detection, i.e., unit edge detection, macro edge detection, and local adjustment MENTO and express edge information in a compact.
[0007] Other purposes of this invention offer the image processing system which fidelity is excellent in in coding of image data, and encodes the image data based on high-pressure shrinking percentage, and its approach. Other purposes of this invention decompose image data into primary image components and secondary image components, and offer the image processing system which can reproduce the primary image by which edge orientation was carried out from primary image components, and its approach.

[Translation done.]

## MEANS

[Means for Solving the Problem] It has the following means, in order to solve the above-mentioned technical problem. While the image processing system concerning invention of the 1st of this application detects the binary image in which the edge location which is the boundary of reception and this image data to a pixel value change about the image data of two-dimensional size is shown The 1st [ which corresponds the 1st edge positional information which shows the direction component of said edge location from this binary image ] detection means which detects for every pixel, The 2nd detection means which detects the 2nd edge positional information which reception and a direction component carry out grouping of said 1st edge positional information relevant to mutual for said 1st edge positional information, and shows the edge location by which grouping was carried out, A means to ask for the transition location of a pixel value by referring to each pixel value in said image data area corresponding to reception and said 2nd edge positional information for said 2nd edge positional information, It has a coding means to encode the difference of the edge location specified to said 2nd edge positional information and said 2nd edge positional information, and said transition location.
[0009] Moreover, a detection means to detect the edge positional information the image processing system concerning the 2nd invention indicates the edge location which is the boundary of reception and this image data to a pixel value about the image data of two-dimensional size to be, A coding means to encode said edge positional information, and an image reconstruction means to perform image reconstruction based on the encoded edge positional information, the difference of said image data and the playback image data reproduced by said playback means — a means to encode a value, the image based on edge positional information, and difference — it has the means and ** which raise image quality gradually by carrying out sequential transmission of the value. Since the application by this invention is a thing relevant to MPEG-4 area, the video telephone which deals with several sorts of bodies, and its pocket communication link are desirable as target application. The test sequence of a bust is used for simulation by such intention.

[Translation done.]

* NOTICES *



---

## OPERATION

---

[Function] According to this invention, by having used hierarchical edge detection (the unit edge detection shown in an example; macro edge detection, and step of local adjustment MENTO), the coding technique of edge contrast directivity can be offered and coding processing of the outstanding image data can be performed in consideration of the moral vision property of human being's vision system.

---

[Translation done.]

EXAMPLE

[Example] Hereafter, the example of the image processing system of this invention is explained to a detail with reference to a drawing. Drawing 1 is the block diagram showing the configuration of decomposition coding concerning the image processing system of this example. As shown in this drawing, a subject-copy image is supplied to the edge extract section 10, and the edge information about the border line of the body in an image etc. is extracted here. The extracted edge information is supplied to the data optimization section 12, and in order to attain high-pressure shrinkage, edge information-redundancy nature is removed. In this way, the primary component (primary component) 14 about the edge information extracted from the subject-copy image is obtained.
[0012] Moreover, the primary component 14 is supplied to the image reconstruction section 16, and the primary image 18 is reproduced based on a primary component. and the reproduced primary image 18 — difference — a subject-copy image and difference ask in a vessel 20 — having — this difference — a value is supplied to the differential-encoding section 22. here — difference — DCT processing by the adjustable size block is performed about a value, and the encoded smooth component (secondary component) 24 is obtained. Drawing 2 is drawing showing the hierarchization edge extract process in the edge extract section of drawing 1 . The process of the edge extract from a subject-copy image is performed through each of the Laplacian filter 100, the unit edge detecting element 102, the macro edge detecting element 104, the local adjustment MENTO section 106, and the on-the-strength count section 108.
[0013] First, as a subject-copy image, the image data (brightness component) of 480*704-pixel size is supplied to the Laplacian filter 100, and well-known Laplacian processing is performed. That is, secondary differential data in which the change about each pixel is shown are called for. In addition, the image data of a color difference component is the size sampled by one half in 480*704, and these data are used in the on-the-strength count section 108 mentioned later. Next, the data by which Laplacian processing was carried out are supplied to the unit edge detecting element 102, and the binary image in which an exact edge location is shown is obtained by using mu+K-sigma for a threshold. Here, mu, sigma, and K are an average, the standard deviation of derivative space, and a multiplier, respectively. The concept of an edge location is used as a thing showing the boundary of the change, when the brightness of each pixel in image data changes steeply and continuously.
[0014] And it matches about the pixel which shows the edge location in a subject-copy image using the segment pattern in which eight directions as shown in drawing 3 are shown. The pattern for matching is shown by Template Tn (n= 0, 1, .. 7), and each entry in (j, k) is expressed by tn (j, k). the sub field which consists lambda (x y) of 5*5pixellambda (x+j, y+k) — it is — making (j, k= 0, 1, 2, 3, 4) . The cross-correlation CRn (x y) between Tn and lambda (x y) is calculated by the degree type.
[Equation 1]

$$CRn\ (x,\ y) = \sum_{j=0}^{4} \sum_{k=0}^{4} tn\ (j,\ k) \times \lambda\ (x+j,\ y+k) \cdots (1)$$

[0015] Therefore, CRn (x y) is equal to 7, or if n which is seven or more exists, a flag will stand

on the coordinate (x y) of n bit plane. This shows that it was chosen as a matching pattern which Template Tn is a coordinate (x y) and calls a unit edge. For example, in matching about the pixel which shows a certain edge location, the pixel concerned used as a processing object is located at the core (location of three-line three trains) of a template, and if the neighboring pixel of the pixel concerned is located horizontally, template T four will be chosen, and the edge location which is the pixel concerned is processed as a thing with the direction component of T four. In addition, among drawing, the numeric value of "1" and "2" shows weighting of matching, and "2" is processing it as a thing heavier than "1" by this example.

[0016] Next, grouping, i.e., macro-izing, is performed by the macro edge detecting element 104 about the pixel by which the unit edge was detected. As mentioned above, although a unit edge is prescribed by the template of eight directions, it is connected to the macro edge on which each of these unit edges are specified in the 16 directions in a continuous phase.

[0017] If the pixel concerned which should be processed shall be located in one line, five trains, and (1, 5) supposing the pixel matrix of five line *9 train, the direction of 16 specified on a macro edge the location of (5, 5) of right under [ the ] — direction "0" — carrying out — and order — (5, 4), and ... (5 3) (5 1) — respectively — a direction — "1" and "2 "and" 4" — carrying out — moreover, order — (4, 1), (3, 1), and (2, 1) — respectively — a direction — it is referred to as "5 "and" 7." the same — the order from direction "0" — right-hand side — (5, 6), and ... (5 7) (5 9) — respectively — a direction — "15" and "14 "and" 12" — carrying out — moreover, order — (4, 9), (3, 9), (2, 9), and (1, 9) — respectively — a direction — it is referred to as "11" and "10 "and" 8." The direction n of a unit edge corresponds to N specified at a ceremony (2), and N is the core of the direction of a base that search processing for connection is performed.

[Equation 2] N= 2n ... (2)

[0018] The direction of the connection in a macro edge must be detected out of the direction of three candidates, i.e., N, N-1, and N+1. For example, if a certain unit edge is a template T1, the unit edge which is N= 2, therefore should be connected from the direction of 1, 2, and 3 will be detected. A unit edge must be connected in the direction which can extract a macro edge with the longest criteria of selection. In the direction of each candidate, it is determined in each node segmented by the unit length Lunit (4 pixels is said) whether a macro edge is connected. If the flag of the bit side n, n-1, and n+1 arises near a node, a macro edge will be extended till a node. It is dependent on what is carried out what magnitude of a field counts for such decision. In addition, the unit edge of the pixel which will be located on the macro edge is eliminated after connection processing termination of a macro edge.

[0019] Here, with reference to drawing 4 and drawing 5 , the concrete example of macro edge detection is explained. A macro edge is detected by connecting a unit edge in the following procedures. First, the unit edge developed on n-bit plane is considered to be seemingly developed by N-bit plane so that it may correspond to (2) types. Supposing the flag stands on N-bit plane now, it will consider as the starting point lambda of the macro edge which detects the location after this. It is node piN and L about the point searched in case L unit edges are connected in the direction N from the starting point lambda. A definition is given and it is lambda in Lunit=4, piN, and L. Physical relationship is shown in drawing 4 . At this example, it is piN and L about a search aperture. It will consider as the 3x3-pixel field made into a core, and if the flag stands on this field on N, N-1, and an N+1-bit plane, sequential increment of the number L of connection is carried out, a search will be repeated and the node of the last whose connection was completed will be made into the terminal point of the candidate macro edge of Direction N. The candidate macro edge of a direction N-1 and a direction N+1 is also detected from the same starting point (if the flag stands into the search aperture on N-2 or N-bit plane by the search of a direction N-1 in these cases, it will be judged as connection), and, finally the number L of connection detects the greatest thing as a macro edge in these 3 candidate. Supposing two or more candidates with the maximum number L of connection exist, this number will choose the larger one, using the total of the flag in a search aperture as secondary scale. Next, the example of connection of a unit edge is shown in drawing 5 . The minimum field divided by the ruled line is a pixel, and each block diagram expresses the same subregion on an image. Let the first location scans N bit plane and the flag stands be the starting point lambda of a macro edge. In this

example, lambda is called for on 2-bit plane. Therefore, the search for connection will search a direction 3 by the search of a direction 1, and 2-bit plane by 0-bit plane at the search of directions 1, 2, and 3, and 4-bit plane. piN in case the number L of connection becomes max in the search of directions 1, 2, and 3, respectively, and L A location is shown in the lower berth of drawing 5 . Finally at this example, it is pi 1 and 3. It becomes the terminal point of the macro edge to detect.

[0020] In this way, detection of the macro edge about each unit edge is performed, and the starting position of the macro edge by which grouping was carried out, the direction of either of 16, and die length are obtained. Next, in the local adjustment MENTO section 106, the rectangle field which wraps a macro edge in predetermined die-length Lext (this example 7 pixels) is specified as an edge belt E using the called-for macro edge. Drawing 6 is the example of a belt edge and sets a shaft perpendicular to a shaft parallel to a macro edge to p and q, respectively. And the pixel value on the edge belt E is expressed as epsilon (x y).

[0021] Generally, an actual edge can be assumed to be what exists along with the macro edge of the edge belt E. In order to locate an actual edge correctly, change (8-bit gradation) of the gray level of a pixel is inspected in accordance with the shaft q perpendicular to a macro edge. The step is explained.

(i) The average phi of the gray level of all the pixels in an edge belt is calculated first.

(ii) If it is smaller than the average psi, 0 is marked on the pixel corresponding to epsilon (x y), and if the gray level of each pixel is larger than the average psi, it will mark 1 about each pixel of the edge belt E.

(iii) An actual edge is located to the place which the transition to 0 to 1, or 1-0 produces (default"0" which expresses the pixel which does not produce such transition about the shaft on a macro edge is used).

(iv) By calculating respectively the average value about the pixel marked by the both sides of 0 and 1, the edge profile which has an ideal step function is approximated, and it is the reinforcement delta 0 of the lower one. Reinforcement delta 1 of the higher one It obtains, respectively.

[Equation 3]

$$\delta_0 = \frac{1}{\tau_0} \sum_{(p,q)} \sum_{<\phi} \varepsilon(p, q) \quad \cdots(3)$$

$$\delta_1 = \frac{1}{\tau_1} \sum_{(p,q)} \sum_{\geq\phi} \varepsilon(p, q) \quad \cdots(4)$$

[0022] Here, it is tau 0. tau 1 The number of pixels with which are satisfied of the monograph affair of a sum total type (3) and (4) is shown. Drawing 7 shows the example of tic [ SUKEMA ] acquired by the local adjustment MENTO section 106. An actual edge location is pursued by the thick wire. Moreover, the sequence of the pixel of a shaft q= 0 corresponds to a macro edge. In addition, the on-the-strength count section 108 performs count of the above-mentioned sum total type (3) and (4) based on the result of the local adjustment MENTO section 106. Moreover, the on-the-strength count section 108 performs count on the strength with the same said of a color difference component.

[0023] The edge data obtained according to the above process are outputted to the data optimization section 12 from the edge extract section 10 shown in drawing 1 . In order to attain high-pressure shrinking percentage, the data optimization section 12 is removed from the edge data from which redundancy and the information which is not not much important were extracted, and encodes edge data.

(i) Since the actual edge obtained by the local adjustment MENTOROKARUAJASUTOMENTO section 106 has the low pass property which met in the direction of a macro edge as shown in above-mentioned drawing 7 , by carrying out subsampling in the predetermined period Lsub, it reduces data and performs differential encoding about the reduced data.

(ii) — reinforcement and Weber FEFUNA — with reference to a brightness difference threshold,

this removes the edge which is not not much important in human being's vision sensibility using law. Let deltaI be the brightness difference threshold of an illuminance I. deltaI is prescribed that brightness becomes remarkable, when a brightness difference reaches deltaI or it is exceeded.
[Equation 4]

$$\frac{\Delta I}{I} = \theta \quad \cdots(5)$$

[0024] The above-mentioned formula can be expressed with a multiplier zeta, being able to assume that it is what deltaI/I gives the magnitude of the vision sensibility E well.
[Equation 5]

$$\Delta E = \zeta \frac{\Delta I}{I} \quad \cdots(6)$$

It integrates with this and is [Equation 6]. E=zeta'logI ... (7)
In order to apply the Weber FEFUNA method to the encoding method, a formula (5) is transformed using an original definition. That is, the macro edge with which are satisfied of a bottom type is removed from the data of a primary component.
[Equation 7] delta1-delta0 <=thetapsi ... (8)
[0025] In an actual case, thetay and thetac can be set about brightness and a chrominance, respectively.
[Table 1]

表 1

| カテゴリ | 符号化されるメッセージ | 範囲 | ビット数 |
|---|---|---|---|
| 幾何学的情報 | 始点 | ピクチャサイズ | $\log_2$（横サイズ）+$\log_2$（縦サイズ） |
| | 方向 | [0, 15] | 4 |
| | Lunit を1単位とした長さ | [1, 32]* | 可変長コード |
| ローカルアジャストメント | 実際のエッジ位置と異なる | [-7, 7]* | 可変長コード |
| 強度 | クロマ重要度フラッグ | 0又は1 | 1（クロミナンスのみ） |
| | ステップ関数のタイプ | 0又は1 | 1 |
| | 低い方の強度：$\delta_0$ | [0, 255] | 6 |
| | コントラスト：$\delta_1-\delta_0$ | [0, 255] | 可変長コード |
| ノート：* 画像サイズ及び/又は手段による | | | |

[0026] Table 1 shows the message for encoding the macro edge about a primary component. As geometric information on a macro edge, the starting point of a macro edge, a direction, and the die length of a unit have the predetermined range and the number of bits, and are encoded. Moreover, about local adjustment MENTO, the difference of a macro edge and an actual edge location is encoded. Moreover, in this example, the element of the both sides of brightness and a chrominance is encoded. Geometric information and geometric local adjustment MENTO are obtained using a luminance element, and reinforcement other than another side and the flag showing the semantics of a chroma (color) is calculated for every color element.
[0027] Next, the image reconstruction section 16 is explained. The primary component 14

contains only the reinforcement of edge associated data, i.e., geometric information, and each macro edge. So, a certain kind of interpolation/extrapolation must be used in order to predict the gray level in fields other than an edge belt. Reconstructive processing draws the pixel in each edge belt using (i) geometry information, local adjustment MENTO, and reinforcement (namely, contrast). About local adjustment MENTO, interpolation of the actual edge location between adjoining sampling points is carried out in linearity. In this way, each pixel in an edge belt can give a gray level depending on the side in which an actual edge is located.
(ii) The reference pixel of eight directions is used for interpolation, omegai which predicts the gray level on pixel criteria, and alphai are made into distance from the pixel predicted to be a reference pixel, and phi on Direction I shows each (refer to drawing 8 ). Other reference pixels do not exist between phi and omega i. omegai is made into the gray level in reference pixel omegai, and the gray level of the pixel which is shown by phi and which is predicted is called for from a bottom type.
[Equation 8]

$$\bar{\phi} = \frac{1}{\sum_{i=0}^{7} \alpha i^{-1}} \sum_{i=0}^{7} \vec{\omega} i \ \alpha i^{-1} \quad \cdots (9)$$

[0028] Such processing is performed until it results far away from a near edge belt, in order to obtain the smooth change by the gray level. The segmentation in an adjustable size block is performed to the whole image, and interpolation is performed from a small block to a big block.
[0029] Next, the differential-encoding section is explained. Although the primary image 18 reproduced by the image reconstruction section 16 is lacked in the detail of a smooth field, i.e., the loose field of a change on the strength, it offers the depiction which was excellent in consciousness. In order to fill up the detail in this field, the so-called adjustable block-size coding processing on the basis of the edge orientation sensibility of human being's vision system is used. That is, coding by the small block size is performed in the neighborhood of an edge, and coding by another side and the big block size is used when the distance from an edge increases. The combination of a block size 8*8, 16*16, and 32*32 is applied in order to attain high-pressure shrinking percentage. In this way, the quality image defined as the second image with the SUKERA kinky thread tee of SNR can be obtained.
[0030] This technique is coding characterized by nonlinear sampling using edge information, and shows the outline of the processing to drawing 9 . it mentioned above — as ( drawing 1 ) — the difference of the subject-copy image 200 and primary component images 18 — a value — difference — it asks with a vessel 20 — having — this difference — a value is supplied to the nonlinear sampling section 202. Nonlinear sampling to which a block size is changed according to the local property of an image is called the adaptation block encoding method, and the various implementation technique is proposed. Those most are transmitting additionally the information which shows a block size, and the information which shows division of a block. In order that this technique may utilize the information on the edge (after local adjustment MENTO application) which is the coded data of a primary component and may realize nonlinear sampling, it does not need additional information. In this example, a square block (the block of three kinds of magnitude, i.e., 32-pixel S*32 base, 16 pixels * 16 pixels, and 8 pixels * 8 pixels) is used. First, an edge is developed on the bit plane of the same magnitude as an image, and a flag is set in the location where the element of an edge exists. Next, a 32-pixel block [ -32 pixel ] performs linearity sampling. And if the flag stands in the block concerned about each block, it divides into four 16 pixel x 16-pixel blocks and the flag does not stand, suppose that it remains as it is. Similarly, in the next phase, if the flag stands in the block concerned about the 16 pixel x 16-pixel block, it divides into four 8 pixel x 8-pixel blocks and the flag does not stand, suppose that it remains as it is. Thus, improvement in vision evaluation is expectable by changing a block size depending on the distance from an edge.
[0031] Processing after nonlinear sampling is performed one by one by the discrete cosine transform section 204 and multiplier quantization which are generally performed and the capable

block judging section 206, the multiplier scan section 208, and the run level coding section 210. It is fundamentally [ as MPEG-1 (Moving Pictures Expert Group Phase-1:ISO/IEC -11172) which is JPEG (Joint Photographic Expert Group;ISO -10918) and the dynamic-image coding standard which are coding using a discrete cosine transform (DCT), for example, a color static-image coding standard, and MPEG-2 (Moving Pictures Expert Group Phase-2:ISO/IEC -13818) ] the same. By the proposal technique, using three kinds (namely, 32 pixel x32 pixel, 16 pixel x16 pixel, and 8 pixels x 8 pixels) of discrete cosine transforms with nonlinear sampling is mentioned to using as difference the discrete cosine transform these criteria of whose are 8 pixel x8 pixels.
[0032] Next, the result of the simulation of this example is shown below. the engine performance of above-mentioned this example — simulating — the intra of MPEG-2 — it compared with the engine performance of image (inside of frame) coding. Simulation conditions are K= 1.0, Lunit=4 pixel, Lext=7 pixel, thetay=0.10, thetac=0.02, and Lsub=4 pixel.
[0033] As shown in drawing 10 , test sequence "Susie" whose pixel size is 704 pixels * 480 lines is used. A color format is 4:2:0 and is 8 bits/pixel. The result of macro edge detection and local adjustment MENTO is shown in drawing 11 and drawing 12 , respectively. a playback image, i.e., a primary image, and the second image — respectively — drawing 13 and 15 — being shown — them and HPEG- SNR to which an image corresponds 2 intra is shown in Table 2.
[Table 2]

表 2

| 符号化法 | ビットレート<br>ビット/フレーム | SNRS(dB) | | |
|---|---|---|---|---|
| | | Y | Cb | Cr |
| 1次画像 | 15, 049 | 22.91 | 36.31 | 34.29 |
| 2次画像 | 57, 168(total) | 32.70 | 41.38 | 40.93 |
| MPEC-2 (I-画像) | 81, 224 | 31.61 | 44.11 | 43.45 |

The segmentation in the adjustable size block used for the both sides of interpolation processing and differential encoding is shown in drawing 14 . In this example, 255 macro edges extracted from the subject-copy image exist. In addition, as for this drawing (a), the 2nd-step division image and this drawing (c) show the last division image, as for an initial division image and this drawing (b).
[0034] The hierarchical edge detection by this example gives the compact expression of edge information, and, so, a primary image gives the rough outline of a body as shown in drawing 13 , or a scene. Although the compressibility about a primary component was 250:1 or more, image quality was not suitable in itself. Addition of a smooth component attains 70:1 or more compressibility, and raises image quality considerably. From the above-mentioned simulation result, the picture compression encoding method by this example is a low bit rate, is SNR which is equal to image coding in MPEG-two frames (intra), and offers high definition more.
[0035] This invention has the description based on the encoding method using the hierarchical edge detection equipped with differential encoding. This approach is accomplished corresponding to the need about the latest low bit rate / high-pressure shrinking percentage image encoding method. The model used in the example of this invention disassembles an image into the primary component containing an edge element, and the smooth component in which a change on the strength carried out slowly is shown fundamentally. The effectiveness of this invention which let pass with simulation and was obtained is as follows.
[0036] MPEG- which mentioned [ 1st ] the technique of this example above — the engine performance of image coding (I-picture) is improved 2 intra, and MSE based on the precision of another side and a body can match. What mainly contributes to the 2nd at such engine-performance amelioration is the effective sensuous tuning [ express and ] on the basis of the

direction and differential encoding, i.e., adjustable block-size coding, of edge information which used hierarchical edge detection. Furthermore, this example enables the gradual transfer to a secondary image from a primary image, it is desirable to application like browsing in an image database etc., and most of another side and second generation coding techniques do not give this description.

---

[Translation done.]

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]
[Drawing 1] The block diagram showing the configuration of decomposition coding concerning the image processing system of this example.
[Drawing 2] Drawing showing the hierarchization edge extract process in the edge extract section of drawing 1 .
[Drawing 3] Drawing showing the segment pattern of eight directions.
[Drawing 4] Drawing showing the physical relationship of lambda and pi in macro edge detection.
[Drawing 5] Drawing showing the example of connection of the unit edge in macro edge detection.
[Drawing 6] Drawing showing the concept of an edge belt.
[Drawing 7] Drawing showing the example of local adjustment MENTO.
[Drawing 8] Drawing showing the relation of phi and omegai in playback of a primary image.
[Drawing 9] Drawing showing the step of the adjustable block-size encoding method.
[Drawing 10] Drawing showing test sequence "Susie".
[Drawing 11] Drawing showing the result of macro edge detection.
[Drawing 12] Drawing showing the result of local adjustment MENTO.
[Drawing 13] Drawing showing a primary image.
[Drawing 14] Drawing showing the segmentation in an adjustable size block.
[Drawing 15] Drawing showing a secondary image.
[Description of Notations]
10 Edge Extract Section
12 Data Optimization Section
14 Primary Component
16 Image Reconstruction Section
18 Primary Image
20 Difference — Vessel
22 Differential-Encoding Section
24 Smooth Component
100 Laplacian Filter
102 Unit Edge Detecting Element
104 Macro Edge Detecting Element
106 Local Adjustment MENTO Section
108 On-the-Strength Count Section

---

[Translation done.]

引用文献1

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平8－126003

(43)公開日　平成8年(1996)5月17日

| (51)Int.Cl.⁶ | 識別記号 | 庁内整理番号 | ＦＩ | 技術表示箇所 |
|---|---|---|---|---|
| Ｈ０４Ｎ　7/24 | | | | |
| Ｇ０６Ｔ　9/00 | | | | |
| Ｈ０４Ｎ　1/41 | B | | | |
| | | | Ｈ０４Ｎ　7/ 13 | Z |
| | | | Ｇ０６Ｆ　15/ 66　３３０ | H |

審査請求　未請求　請求項の数6　ＯＬ　（全 11 頁）　最終頁に続く

(21)出願番号　特願平6－240325

(22)出願日　平成6年(1994)10月4日

(71)出願人　590000879
テキサス　インスツルメンツ　インコーポレイテツド
アメリカ合衆国テキサス州ダラス，ノース　セントラルエクスプレスウエイ　13500

(72)発明者　伊藤　裕二
茨城県つくば市御幸が丘17番地　テキサス・インスツルメンツ筑波研究開発センター内

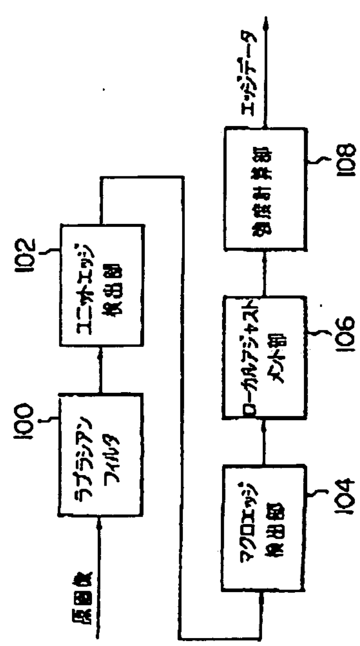
(74)代理人　弁理士　浅村　皓　（外3名）

(54)【発明の名称】　画像処理装置及びその方法

(57)【要約】

【目的】　本発明の目的は、階層的エッジ検出、すなわちユニットエッジ検出、マクロエッジ検出、及びローカルアジャストメントのステップにより、画像データからエッジ位置情報を得ることができる画像処理装置を提供する。

【構成】　本発明に係る画像処理装置は、２次元サイズの画像データを受け取り、エッジ位置を示す２値画像を検出（１００）し、エッジ位置の方向成分を示すユニットエッジ位置情報を検出するユニットエッジ検出部（１０２）と、ユニットエッジ位置情報の方向成分が相互に関連するものをグループ化し、マクロエッジ位置情報を検出するマクロエッジ検出部（１０３）と、マクロエッジ位置情報を受け取り、マクロエッジ位置情報に対応する画像データ領域内の各画素値を参照することにより画素値の遷移位置を求めるローカルアジャストメント部（１０６）と、マクロエッジ位置情報と、マクロエッジ位置及び遷移位置との差分とを符号化する符号化部とを備える。



【特許請求の範囲】

【請求項1】 （ａ）２次元サイズの画像データを受け取り、該画像データから画素値の変化の境界であるエッジ位置を示す２値画像を検出するとともに、該２値画像から前記エッジ位置の方向成分を示す第１のエッジ位置情報を対応する画素毎に検出する第１の検出手段と、（ｂ）前記第１のエッジ位置情報を受け取り、方向成分が相互に関連する前記第１のエッジ位置情報をグループ化し、グループ化されたエッジ位置を示す第２のエッジ位置情報を検出する第２の検出手段と、（ｃ）前記第２のエッジ位置情報を受け取り、前記第２のエッジ位置情報に対応する前記画像データ領域内の各画素値を参照することにより画素値の遷移位置を求める手段と、（ｄ）前記第２のエッジ位置情報と、前記第２のエッジ位置情報に規定されるエッジ位置と前記遷移位置との差分とを符号化する符号化手段とを有する画像処理装置。

【請求項2】 請求項第１項において、前記第２のエッジ位置情報には、グループ化されたエッジの長さ、方向、及び開始位置を示す幾何学的情報が含まれることを特徴とする画像処理装置。

【請求項3】 請求項第１項において、前記画素値の遷移位置を求める手段は、前記画像データ領域内の各画素値の平均値と、各画素値とを比較することにより行われることを特徴とする画像処理装置。

【請求項4】 （ａ）２次元サイズの画像データを受け取り、該画像データから画素値の境界であるエッジ位置を示すエッジ位置情報を検出する検出手段と、（ｂ）前記エッジ位置情報を符号化する符号化手段と、（ｃ）符号化されたエッジ位置情報に基づき画像再生を行う画像再生手段と、（ｄ）前記画像データと前記再生手段により再生された再生画像データとの差分値を符号化する手段と、（ｅ）エッジ位置情報に基づく画像と差分値を順次伝送することにより段階的に画質を向上させる手段と、を有する画像処理装置

【請求項5】 請求項第４項において、前記符号化手段は、エッジ位置情報を利用した非線形標本化により前記差分値を符号化する手段を含むことを特徴とする画像処理装置。

【請求項6】 （ａ）２次元サイズの画像データを受け取り、該画像データから画素値の変化の境界であるエッジ位置を示す２値画像を検出し、（ｂ）前記２値画像から前記エッジ位置の方向成分を示す第１のエッジ位置情報を対応する画素毎に検出し、（ｃ）前記第１のエッジ位置情報を受け取り、方向成分が相互に関連する前記第１のエッジ位置情報をグループ化し、グループ化されたエッジ位置を示す第２のエッジ位置情報を検出し、（ｄ）前記第２のエッジ位置情報を受け取り、前記第２のエッジ位置情報に対応する前記画像データ領域内の各画素値を参照することにより画素値の遷移位置を求め、（ｅ）前記第２のエッジ位置情報と、前記第２のエッジ位置情報に規定されるエッジ位置と前記遷移位置との差分とを符号化するステップを有する画像処理方法。

【発明の詳細な説明】

【０００１】

【従来の技術】国際標準化機構（ＩＳＯ）は、低ビット速度で動画を符号化する一般的な方法の必要性に応答して、ビデオ電話や移動体通信などの種々のアプリケーションについてのビデオ符号化標準を構築するため、ＭＰＥＧ－４（Ｍｏｖｉｎｇ Ｐｉｃｔｕｒｅｓ Ｅｘｐｅｒｔ Ｇｒｏｕｐ，Ｐｈａｓｅ ４）を１９９３年に形成した。多くの画像圧縮符号化法の主な目的は、原画像についての忠実度の高い再生をできるだけ高圧縮率で行なおうとする。圧縮符号化法の設計に用いられる忠実度の基準は、性能上大きな役割を果たす。一般に用いられている忠実度の基準は、平均自乗誤差（ＭＳＥ）である。ＭＳＥの主な特質は、その数学的な計算が容易であることと、ＭＳＥの小さな値が感覚的に高品質の再生画像に実際に対応することである。後者は、人間の目によって、再生された画像の最終的な判断が行われるため重要である。

【０００２】忠実度の基準としてＭＳＥを有する、変換符号化などの種々の画像符号化技術が開発されてきた。これらの技術は、１．０ビット／画素か、それ以上のビット速度で比較的高品質な再生画像をもたらすが、その一方で、特殊な可視的な劣化、例えば、ブロック歪、エッジボケなどを、しばしば、より低ビット速度で生じさせる。過去１０年以上にわたり、第２世代符号化技術としてよく知られている、画像符号化法の新しいクラスが開発されてきた。これらの方法は、非常に低速のビット速度でＭＳＥ指向の手法の品質向上を証明し、輪郭線やテキスチャのような真の実体を一層コンパクトに表現することによって画像を描こうとするものである。それ故、第２世代符号化法は、未だ実現されていない人間の目に指向された忠実度の基準で高圧縮率を達成することを期待されている。

【０００３】スケッチ画に基づく画像符号化は、このカテゴリーの典型的な手段として考えられており、参考として引用されてきた。この方法は、輪郭線抽出プロセスにより特徴付けを行うものであり、画素単位の検出により局所的なノイズに対して脆弱であるため、スケッチ画再生において、有害な歪をもたらす結果となる。

【０００４】

【発明が解決しようとする課題】また、上述の手法以外に、従来より原画像を２つの構成要素に分解して、それらの構成要素を符号化する手法が知られている。この手法は、原画像を、エッジ情報を含むプライマリコンポーネント（１次成分）と、輝度変化の緩やかなスムースコンポーネント（２次成分）とに分解する。分解符号化の考えは、ヤン（Ｙａｎ）及びサクリソン（Ｓａｋｒｉｓｏｎ）による２成分モデルから生み出され、それ以来、

いくつかの実践的なアプローチが研究されてきた。

【0005】これらの分解符号化によるすべてのアプローチは、エッジ抽出、すなわち、検出、表現及び符号化を中心とするものであり、これは、再生画像の画質、ひいては全体の性能に大きなインパクトを有する。大部分の手法は、ピーク点追跡、つまりエッジ追跡のような画素単位のエッジ検出を行い、かつ、エッジ位置のシーケンスを符号化するためにチェーン符号化を用いる。しかしながら、上記手法は、閉曲線の問題や、局所的なノイズへの対処が困難であり、エッジの中断、位置エラー及び不正確な強度の結果を招く、といったことが指摘されている。また、前後関係の情報は、非常に限られた領域にしか存在しないと考えられるため、物体の境界に対応する輪郭線と、そうでない輪郭線とを区別する方法はなかった。

【0006】本発明の目的は、上記従来技術の課題を解決し、階層的エッジ検出により、画像データからエッジ位置情報を得ることができる画像処理装置及びその方法を提供する。本発明の他の目的は、階層的エッジ検出、すなわちユニットエッジ検出、マクロエッジ検出、及びローカルアジャストメントの各ステップを導入し、エッジ情報をコンパクトに表現するとともに、高品質の画像を提供する。

【0007】本発明の他の目的は、画像データの符号化を、忠実度が優れ、かつ高圧縮率による画像データの符号化を行う画像処理装置及びその方法を提供する。本発明の他の目的は、画像データを1次画像成分及び2次画像成分に分解し、1次画像成分からエツジ指向された1次画像を再生することができる画像処理装置及びその方法を提供する。

【0008】

【課題を解決するための手段】上記課題を解決するために、以下の手段を備えるものである。本願の第1の発明に係る画像処理装置は、2次元サイズの画像データを受け取り、該画像データから画素値の変化の境界であるエッジ位置を示す2値画像を検出するとともに、該2値画像から前記エッジ位置の方向成分を示す第1のエッジ位置情報を対応する画素毎に検出する第1の検出手段と、前記第1のエッジ位置情報を受け取り、方向成分が相互に関連する前記第1のエッジ位置情報をグループ化し、グループ化されたエッジ位置を示す第2のエッジ位置情報を検出する第2の検出手段と、前記第2のエッジ位置情報を受け取り、前記第2のエッジ位置情報に対応する前記画像データ領域内の各画素値を参照することにより画素値の遷移位置を求める手段と、前記第2のエッジ位置情報と、前記第2のエッジ位置情報に規定されるエッジ位置と前記遷移位置との差分とを符号化する符号化手段とを有する。

【0009】また、第2の発明に係る画像処理装置は、2次元サイズの画像データを受け取り、該画像データから画素値の境界であるエッジ位置を示すエッジ位置情報を検出する検出手段と、前記エッジ位置情報を符号化する符号化手段と、符号化されたエッジ位置情報に基づき画像再生を行う画像再生手段と、前記画像データと前記再生手段により再生された再生画像データとの差分値を符号化する手段と、エッジ位置情報に基づく画像と差分値を順次伝送することにより段階的に画質を向上させる手段ととを有する。本発明によるアプリケーションは、MPEG－4エリアに関連するものであるため、ターゲットアプリケーションとして、数種の物体を取り扱うビデオ電話や携帯通信が好ましい。こうした意向により、胸像のテストシーケンスがシミュレーションに用いられる。

【0010】

【作用】本発明によれば、階層的エッジ検出（実施例に示すユニットエッジ検出、マクロエッジ検出及びローカルアジャストメントのステップ）を用いたことにより、エッジコントラスト指向性の符号化手法を提供することができ、人間の視覚システムの精神視覚性質を考慮し、優れた画像データの符号化処理を行うことができる。

【0011】

【実施例】以下、本発明の画像処理装置の実施例について図面を参照して詳細に説明する。図1は、本実施例の画像処理装置に係る分解符号化の構成を示すブロック図である。同図に示すように、原画像は、エッジ抽出部10に供給され、ここで画像内の物体の輪郭線などに関するエッジ情報が抽出される。抽出されたエッジ情報は、データ最適化部12に供給され、高圧縮化を図るためにエッジ情報の冗長性が取り除かれる。こうして、原画像から抽出されたエッジ情報に関するプライマリコンポーネント（1次成分）14が得られる。

【0012】また、プライマリコンポーネント14は、画像再生部16へ供給され、プライマリコンポーネントに基づき1次画像18が再生される。そして、再生された1次画像18は、差分器20において、原画像と差分が求められ、この差分値は、差分符号化部22へ供給される。ここで、差分値について可変サイズブロックによるDCT処理が行われ、符号化されたスムースコンポーネント（2次成分）24が得られる。図2は、図1のエッジ抽出部における階層化エッジ抽出プロセスを示す図である。原画像からのエッジ抽出のプロセスは、ラプラシアンフィルタ100、ユニットエッジ検出部102、マクロエッジ検出部104、ローカルアジャストメント部106、及び強度計算部108の各々を介して行われる。

【0013】先ず、原画像として、480＊704画素サイズの画像データ（輝度成分）が、ラプラシアンフィルタ100へ供給され、周知のラプラシアン処理が施される。つまり、各画素についての変化を示す2次微分データが求められる。なお、色差成分の画像データは、4

80＊704を1／2でサンプリングされたサイズであり、これらのデータは、後述する強度計算部108において用いられる。次に、ラプラシアン処理されたデータは、ユニットエッジ検出部102に供給され、正確なエッジ位置を示す2値画像が、$\mu+K\cdot\sigma$をしきい値に用いることで得られる。ここで、$\mu$、$\sigma$、Kは、それぞれ平均、導関数空間の標準偏差、及び係数である。エッジ位置の概念は、画像データ内の各画素の輝度が急峻に、かつ連続的に変化する場合、その変化の境界を表すものとして用いられる。

【0014】そして、図3に示すような、8つの方向を＊示すセグメントパターンを用い、原画像内のエッジ位置を示す画素についてマッチングを行う。マッチングのためのパターンは、テンプレイトTn（n＝0，1...7）によって示され、（j，k）での各エントリは、tn（j，k）によって表される。Λ（x，y）を、5＊5画素λ（x＋j，y＋k）からなるサブ領域であるようにする（j，k＝0，1，2，3，4）。TnとΛ（x，y）間の相互相関CRn（x，y）は、次式によって計算される。

【数1】

$$CRn(x,y)=\sum_{j=0}^{4}\sum_{k=0}^{4}tn(j,k)\times\lambda(x+j,y+k)\cdots(1)$$

【0015】従って、もし、CRn（x，y）が7に等しいか、あるいは、7以上であるようなnが存在すれば、フラグが、nビット平面の座標（x，y）に立つ。これは、テンプレイトTnが座標（x，y）で、ユニットエッジと呼ぶマッチングパターンとして選択されたことを示す。例えば、あるエッジ位置を示す画素についてのマッチングにおいて、処理対象となる当該画素を、テンプレイトの中心（3行3列の位置）に位置させ、仮に、当該画素の近隣画素が水平方向に位置しているのであれば、テンプレイトT4が選択され、当該画素のエッジ位置は、T4の方向成分を持つものとして処理される。なお、図中”1”及び”2”の数値は、マッチングの重み付けを示すものであり、本実施例では、”2”が”1”よりも重いものとして処理している。

【0016】次に、マクロエッジ検出部104により、ユニットエッジの検出された画素についてグループ化、すなわち、マクロ化が行われる。上述のように、ユニットエッジは8つの方向のテンプレイトによって規定されるが、これらの各ユニットエッジを、連続的な相において、16方向に規定されるマクロエッジに接続する。

【0017】マクロエッジに規定される16の方向は、5行＊9列の画素マトリックスを想定し、処理すべき当該画素が1行、5列、（1，5）に位置するものとすると、その真下の（5，5）の位置を方向”0”とし、それから順に（5，4）、（5，3）・・・（5，1）をそれぞれ方向”1”、”2”・・”4”とし、また、順に（4，1）、（3，1）、（2，1）をそれぞれ方向”5”、・・・”7”とする。同様に、方向”0”から順に右側に（5，6）、（5，7）・・・（5，9）をそれぞれ方向”15”、”14”・・・”12”とし、また、順に（4，9）、（3，9）、（2，9）、（1，9）をそれぞれ方向”11”、”10”・・・”8”とする。ユニットエッジの方向nは、式（2）に規定されるNに対応し、Nは、接続のためのサーチ処理が行われる基本方向の中心である。

【数2】N＝2n・・・（2）

【0018】マクロエッジにおける接続の方向は、3つの候補、すなわち、N，N−1，N＋1の方向の中から検出されなければならない。例えば、あるユニットエッジがテンプレイトT1であれば、N＝2であり、従って、1、2、及び3の方向から接続されるべきユニットエッジが検出される。選択の基準は、一番長いマクロエッジを抽出することができる方向に、ユニットエッジが接続されなければならない。各候補の方向において、マクロエッジが接続されるかどうかは、ユニット長Lunit（4画素をいう）によってセグメント化された各接続点で決定される。もし、ビット面n，n−1，n＋1のフラグが接続点付近に生じるならば、マクロエッジが、接続点まで延長される。どの程度の大きさの領域が、こうした決定のためにカウントされるかは、実施するものに依存する。なお、マクロエッジの接続処理終了後、そのマクロエッジ上に位置することとなる画素のユニットエッジは消去される。

【0019】ここで、図4及び図5を参照して具体的なマクロエッジ検出例を説明する。ユニットエッジを以下の手順で接続することによりマクロエッジを検出する。まず、n−ビット平面上に展開されているユニットエッジを（2）式に対応するように見かけ上N−ビット平面に展開されていると考える。いま、N−ビット平面上にフラグが立っているとすると、その位置をこれから検出するマクロエッジの始点Λとする。始点Λから方向Nにユニットエッジを L個接続する際にサーチする点を接続点$\Pi_{n,L}$と定義し、$L_{unit}$＝4の場合のΛと$\Pi_{n,L}$との位置関係を図4に示す。この実施例ではサーチ窓を$\Pi_{n,L}$を中心とする3×3画素領域としており、N，N−1，N＋1−ビット平面上でこの領域にフラグが立っていれば接続数Lを順次インクリメントしてサーチを繰り返し、接続ができた最後の接続点を方向Nの候補マクロエッジの終点とする。同一始点から方向N−1及び方向N＋1の候補マクロエッジも検出し（これらの場合、例えば方向N−1のサーチではN−2もしくはN−ビット平面上でサーチ窓の中にフラグが立っていれば接続と判

断する）、これら3候補の中で接続数Lが最大のものを最終的にマクロエッジとして検出する。もし、最大の接続数Lを持つ候補が複数存在するならば、2次の尺度としてサーチ窓中のフラグの総数を用い、この数が大きい方を選択する。次に、ユニットエッジの接続例を図5に示す。罫線で区切られた最小の領域は画素であり、各構成図は画像上の同一の部分領域を表している。Nビット平面を走査してフラグが立っている最初の位置をマクロエッジの始点Aとする。この例では、2－ビット平面上でAが求められる。従って、接続のためのサーチは、0－ビット平面で方向1のサーチ、2－ビット平面で方向1、2及び3のサーチ、4－ビット平面で方向3のサーチをすることになる。方向1、2及び3のサーチにおいて、それぞれ接続数Lが最大になる場合の$\Pi_{\ldots}$の位置を図5の下段に示す。この例では、最終的に$\Pi_{3,,}$が検出するマクロエッジの終点となる。

【0020】こうして、各ユニットエッジについてのマクロエッジの検出が行われ、グループ化されたマクロエッジの開始位置、16のいずれかの方向、及び長さが得られる。次に、ローカルアジャストメント部106において、求められたマクロエッジを用い、所定の長さLext（本実施例では、7画素）でマクロエッジを包む矩形領域をエッジベルトEとして規定する。図6は、ベルトエッジの例であり、マクロエッジに平行な軸と垂直な軸をそれぞれp，qとする。そして、エッジベルトE上の画素値を、ε（x，y）として表す。

【0021】一般に、実際のエッジは、エッジベルトEのマクロエッジに沿って存在するものと仮定することができる。実際のエッジを正確に位置させるために、画素のグレイレベルの変化（8ビット階調）を、マクロエッジに垂直な軸qに沿って検査する。そのステップを説明する。

（i）先ず、エッジベルト内のすべての画素のグレイレベルの平均値φを計算する。

（ii）もし、各画素のグレイレベルが、平均値φよりも小さいならば、ε（x，y）に対応する画素に0をマークし、平均値φよりも大きければ、エッジベルトEの各画素について1をマークする。

（iii）0から1、または1から0への遷移が生じる所へ実際のエッジを位置させる（マクロエッジ上の軸について、そのような遷移を生じない画素を表すデフォルト値”0”を使用する）。

（iv）0と1の双方によってマークされた画素についての平均値を各々計算することにより、理想的なステップ関数を有するエッジプロファイルを近似し、低い方の強度$\delta_0$と、高い方の強度$\delta_1$をそれぞれ得る。

【数3】

$$\delta_0=\frac{1}{\tau_0}\sum_{(p,q)}\sum_{<\phi}\varepsilon(p,q)\qquad\cdots(3)$$

$$\delta_1=\frac{1}{\tau_1}\sum_{(p,q)}\sum_{\geq\phi}\varepsilon(p,q)\qquad\cdots(4)$$

【0022】ここで、$\tau_0$と$\tau_1$は、合計式（3）、（4）の各条件を満足する画素数を示す。図7は、ローカルアジャストメント部106により得られたスケマチック例を示すものである。実際のエッジ位置は、太線で追跡される。また、軸q＝0の画素のシーケンスがマクロエッジに対応する。なお、強度計算部108は、ローカルアジャストメント部106の結果に基づき、上記合計式（3）、（4）の計算を行う。また、強度計算部108は、色差成分についても同様の強度計算を行う。

【0023】以上のプロセスによって得られたエッジデータは、図1に示すエッジ抽出部10からデータ最適化部12へ出力される。データ最適化部12は、高圧縮率を達成するために、冗長性及びあまり重要でない情報を抽出されたエッジデータから除去し、エッジデータを符号化する。

（i）ローカルアジャストメント

ローカルアジャストメント部106によって得られた実際のエッジは、上述の図7に示すように、マクロエッジ方向に沿った低域通過特性を有するため、所定期間Lsubでサブサンプリングすることによりデータを削減し、削減されたデータについて差分符号化を行う。

（ii）強度

また、ウエバーフェフナ法を用いて、輝度差しきい値を参照し、これによって、人間の視覚感度においてあまり重要でないエッジを除去する。ΔIを照度Iの輝度差しきい値とする。ΔIは、輝度差がΔIに到達するか、または、それを越えるときに、輝度が顕著になるように規定される。

【数4】

$$\frac{\Delta I}{I}=\theta\qquad\cdots(5)$$

【0024】ΔI／Iが視覚感度Eの大きさを良く与えるものと仮定して、上記式は、係数ζで表すことができる。

【数5】

$$\Delta E=\zeta\frac{\Delta I}{I}\qquad\cdots(6)$$

これを積分して、

【数6】$E=\zeta'\log I\cdots(7)$

ウエバーフェフナ法を符号化法に適用するために、独自の定義を用いて式（5）を変形する。すなわち、下式を満足するマクロエッジが、プライマリコンポーネントの

データから除去される。

【数7】$\delta_1-\delta_0 \leq \theta\psi$・・・(8)

【0025】実際のケースでは、輝度及びクロミナンス＊について、それぞれθyとθcをセットすることができる。

【表1】

表 1

| カテゴリ | 符号化されるメッセージ | 範囲 | ビット数 |
|---|---|---|---|
| 幾何学的情報 | 始点 | ピクチャサイズ | $\log_2$（横サイズ）$+\log_2$（縦サイズ） |
| | 方向 | [0,15] | 4 |
| | Lunit を1単位とした長さ | [1,32]* | 可変長コード |
| ローカルアジャストメント | 実際のエッジ位置と異なる | [-7,7]* | 可変長コード |
| 強度 | クロマ重要度フラッグ | 0又は1 | 1（クロミナンスのみ） |
| | ステップ関数のタイプ | 0又は1 | 1 |
| | 低い方の強度：$\delta_0$ | [0,255] | 6 |
| | コントラスト：$\delta_1-\delta_0$ | [0,255] | 可変長コード |
| ノート：* 画像サイズ及び／又は手段による | | | |

【0026】表1は、プライマリコンポーネントに関するマクロエッジを符号化するためのメッセージを示すものである。マクロエッジの幾何学的な情報として、マクロエッジの始点、方向、及びユニットの長さが、所定の範囲、及びビット数をもって符号化される。また、ローカルアジャストメントについては、マクロエッジと実際のエッジ位置との差が符号化される。また、本実施例では、輝度及びクロミナンスの双方の要素を符号化する。幾何学的な情報及びローカルアジャストメントは、ルミナンス要素を用いて得られ、他方、クロマ（色）の意味を表すフラグ以外の強度は、各カラー要素ごとに計算される。

【0027】次に、画像再生部16について説明する。プライマリコンポーネント14は、エッジ関連データ、すなわち、幾何情報及び各マクロエッジの強度のみを含む。それ故、ある種の内挿／外挿法が、エッジベルト以外の領域内のグレイレベルを予測するために用いられなければならない。再生プロセスは、

（ｉ）幾何情報、ローカルアジャストメント及び強度（すなわち、コントラスト）を用い、各エッジベルト内の画素を描く。ローカルアジャストメントに関しては、隣接するサンプリング点間の実際のエッジ位置が線形的に内挿される。こうして、エッジベルト内の各画素は、実際のエッジが位置する側に依存してグレイレベルを与えられる。

（ｉｉ）8方向の参照画素を内挿に用いて、画素基準上のグレイレベルを予測するωｉとαｉを、参照画素と予測される画素からの距離にし、それぞれを、方向ｉ上のφによって示す（図8参照）。他の参照画素は、φとωｉ間に存在しない。ωｉを参照画素ωｉでのグレイレベルにし、そして、φによって示される予測される画素のグレイレベルが、下式から求められる。

【数8】

$$\bar{\phi} = \frac{1}{\sum_{i=0}^{7} \alpha i^{-1}} \sum_{i=0}^{7} \bar{\omega} i \quad \alpha i^{-1} \qquad \cdots(9)$$

【0028】グレイレベルでの滑らかな変化を得るために、近接のエッジベルトから遠方に至るまで、このような処理が行われる。可変サイズブロックでのセグメンテーションが、画像全体に施され、そして、小さなブロックから大きなブロックまで内挿が行われる。

【0029】次に、差分符号化部について説明する。画像再生部16によって再生された1次画像18は、なめらかな領域、つまり、強度変化の緩やかな領域での詳細を欠如しているが、知覚的に優れた描写を提供する。この領域での詳細を補充するために、人間の視覚システムのエッジ指向感度を基礎とする、いわゆる可変ブロックサイズ符号化処理が用いられる。すなわち、小さなブロックサイズによる符号化はエッジの近傍で行われ、他

方、大きなブロックサイズによる符号化はエッジからの距離が増加する時に使用される。ブロックサイズ８＊８、１６＊１６及び３２＊３２の組み合わせが、高圧縮率を達成するために適用される。こうして、ＳＮＲのスケラビリティによって第２次画像として定義される高品質な画像を得ることができる。

【００３０】本手法は、エッジ情報を利用した非線形標本化を特徴とする符号化であり、その処理の概要を図９に示す。上述したように（図１）、原画像２００と１次成分画像１８との差分値が差分器２０によって求められ、この差分値が非線形標本化部２０２へ供給される。画像の局所的な性質に応じてブロックサイズを変化させる非線形標本化は適応ブロック符号化法と呼ばれ、様々な実現手法が提案されている。それらの大部分は、ブロックサイズを示す情報やブロックの分割を示す情報を付加的に伝送している。本手法は１次成分の符号化データであるエッジ（ローカルアジャストメント適用後）の情報を活用して非線形標本化を実現するため付加的な情報を必要としない。本実施例では、３種類の大きさのブロック、すなわち３２画素Ｓ＊３２素、１６画素＊１６画素および８画素＊８画素の正方形ブロック、を用いる。まず、画像と同じ大きさのビット平面上にエッジを展開し、エッジの要素が存在する位置にフラグを立てる。次に、３２画素－３２画素ブロックで線形標本化を行なう。そして、各ブロックについて当該ブロック内にフラグが立っていれば４つの１６画素×１６画素ブロックに分割し、フラグが立っていなければそのままとする。同様に次の段階では１６画素×１６画素ブロックについて当該ブロック内にフラグが立っていれば４つの８画素×８画素ブロックに分割し、フラグが立っていなければそのままとする。このように、エッジからの距離に依存してブロックサイズを変化させることにより、視覚評価の向上が期待できる。＊

＊【００３１】非線形標本化以降の処理は、一般的に行われている離散コサイン変換部２０４、係数量子化及び有為ブロック判定部２０６、係数走査部２０８及びラン・レベル符号化部２１０によって順次行われる。離散コサイン変換（ＤＣＴ）を用いた符号化、例えばカラー静止画像符号化標準であるＪＰＥＧ（Ｊｏｉｎｔ Ｐｈｏｔｏｇｒａｐｈｉｃ Ｅｘｐｅｒｔ Ｇｒｏｕｐ；ＩＳＯ－１０９１８）や動画像符号化標準であるＭＰＥＧ－１（Ｍｏｖｉｎｇ Ｐｉｃｔｕｒｅｓ Ｅｘｐｅｒｔ Ｇｒｏｕｐ Ｐｈａｓｅ－１：ＩＳＯ／ＩＥＣ－１１１７２）及びＭＰＥＧ－２（Ｍｏｖｉｎｇ Ｐｉｃｔｕｒｅｓ Ｅｘｐｅｒｔ Ｇｒｏｕｐ Ｐｈａｓｅ－２：ＩＳＯ／ＩＥＣ－１３８１８）と基本的に同じである。相違点としては、これらの標準が８画素×８画素の離散コサイン変換を用いているのに対し、提案手法では非線形標本化に伴い３種類（すなわち、３２画素×３２画素、１６画素×１６画素および８画素×８画素）の離散コサイン変換を用いていることが挙げられる。

【００３２】次に、本実施例のシミュレーションの結果を以下に示す。上記本実施例の性能をシミュレートし、ＭＰＥＧ－２のイントラ画像（フレーム内）符号化の性能と比較した。シミュレーション条件は、Ｋ＝１．０、Ｌｕｎｉｔ＝４画素、Ｌｅｘｔ＝７画素、θｙ＝０．１０、θｃ＝０．０２、Ｌｓｕｂ＝４画素である。

【００３３】図１０に示すように、画素サイズが７０４画素＊４８０行のテストシーケンス"Ｓｕｓｉｅ"を用いる。カラーフォーマットは、４：２：０であり、また、８ビット／画素である。マクロエッジ検出及びローカルアジャストメントの結果を、それぞれ図１１及び図１２に示す。再生画像、すなわち１次画像及び第２次画像をそれぞれ図１３、１５に示し、それら及びＨＰＥＧ－２イントラ画像の対応するＳＮＲを表２に示す。

【表２】

表 2

| 符号化法 | ビットレート<br>ビット/フレーム | SNRS(dB) | | |
|---|---|---|---|---|
| | | Y | Cb | Cr |
| 1次画像 | 16,049 | 22.81 | 36.31 | 34.29 |
| 2次画像 | 57,168(total) | 32.70 | 41.38 | 40.93 |
| MPEG−2（I−画像） | 81,224 | 31.61 | 44.11 | 43.45 |

内挿処理及び差分符号化の双方に用いられる可変サイズブロックでのセグメンテーションを図１４に示す。この例では、原画像から抽出されたマクロエッジは２５５個存在する。なお、同図（ａ）は初期分割画像、同図（ｂ）は第２段階分割画像、同図（ｃ）は最終分割画像を示す。

【００３４】本実施例による階層的エッジ検出は、エッジ情報のコンパクトな表現を与えるものであり、それ故、１次画像は、図１３に示されるような物体やシーンの大まかな概要を与える。プライマリコンポーネントに関する圧縮率は、２５０：１以上であるが、画質は、それ自身では適切なものではなかった。スムースコンポー

ネントの追加は、圧縮率70：1以上を達成し、かなり画質を向上させる。上記シミュレーション結果から、本実施例による画像圧縮符号化法は、低ビット速度で、MPEG－2フレーム内（イントラ）画像符号化と匹敵するSNRで、より高画質を提供する。

【0035】本発明は、差分符号化を備えた階層的エッジ検出を用いた符号化法に基づく特徴を有するものである。このアプローチは、最近の低ビット速度／高圧縮率画像符号化法についての必要性に対応して成されたものである。本発明の実施例で用いられたモデルは、基本的に、エッジ要素を含むプライマリコンポーネントと、ゆっくりとした強度変化を示すスムースコンポーネントとに画像を分解する。シミュレーションと通して得た本発明の効果は、以下のようである。

【0036】第1に、本実施例の手法は、上述したようなMPEG－2イントラ画像符号化（I－ピクチャー）の性能を改良し、他方、物体の精度に基づくMSEは匹敵できる。第2に、このような性能改良に主に貢献するものは、階層的エッジ検出を用いたエッジ情報の効果的な表し方と、差分符号化、つまり可変ブロックサイズ符号化を基礎とした感覚的なチューニングである。さらに、本実施例は、1次画像から2次画像への段階的転送を可能にし、画像データベースなどにおけるブラウジングのようなアプリケーションに好ましく、他方、第2世代符号化技術の大部分はこの特徴を与えるものではない。

【0037】

【発明の効果】本発明に係る画像処理装置によれば、階層エッジ検出（実施例に示すユニットエッジ検出、マクロエッジ検出及びローカルアジャストメントのステップ）を用いたことにより、エッジコントラスト指向性の符号化法を提供することができ、また、人間の視覚システムの精神視覚性質を考慮し、優れた画像データの符号化処理を行うことができる。

【図面の簡単な説明】

【図1】本実施例の画像処理装置に係る分解符号化の構成を示すブロック図。

【図2】図1のエッジ抽出部における階層化エッジ抽出プロセスを示す図。

【図3】8方向のセグメントパターンを示す図。

【図4】マクロエッジ検出におけるΛとΠとの位置関係を示す図。

【図5】マクロエッジ検出におけるユニットエッジの接続例を示す図。

【図6】エッジベルトの概念を示す図。

【図7】ローカルアジャストメントの例を示す図。

【図8】1次画像の再生におけるφとωiとの関係を示す図。

【図9】可変ブロックサイズ符号化法のステップを示す図。

【図10】テストシーケンス”Susie”を示す図。

【図11】マクロエッジ検出の結果を示す図。

【図12】ローカルアジャストメントの結果を示す図。

【図13】1次画像を示す図。

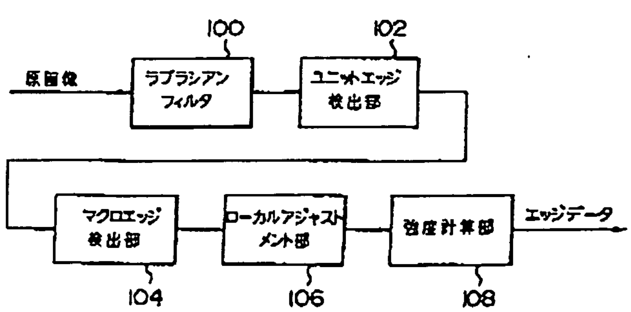
【図14】可変サイズブロックでのセグメンテーションを示す図。

【図15】2次画像を示す図。

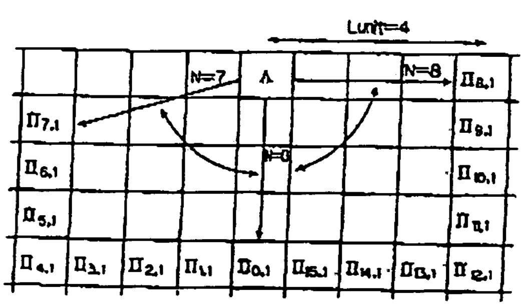
【符号の説明】

10 エッジ抽出部
12 データ最適化部
14 プライマリコンポーネント
16 画像再生部
18 1次画像
20 差分器
22 差分符号化部
24 スムースコンポーネント
100 ラプラシアンフィルタ
102 ユニットエッジ検出部
104 マクロエッジ検出部
106 ローカルアジャストメント部
108 強度計算部

【図2】

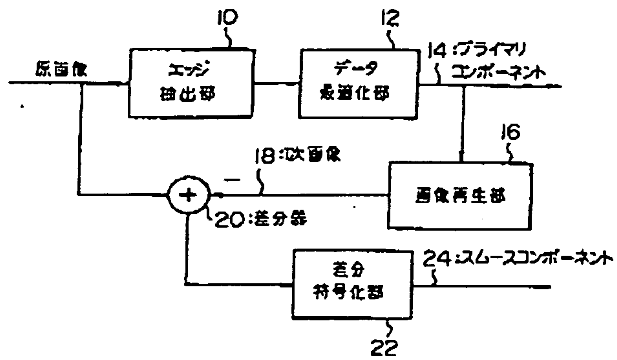



【図4】

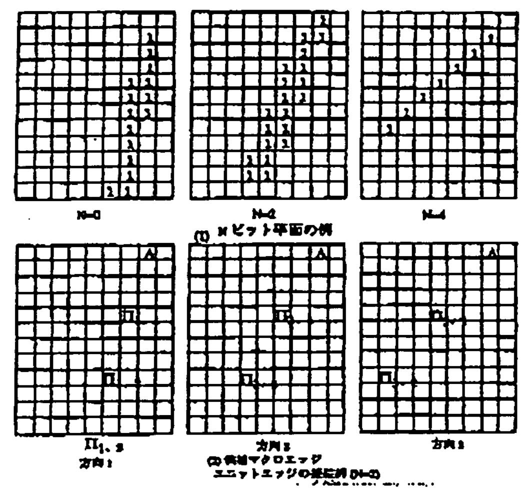



ΛとΠn,1との位置関係

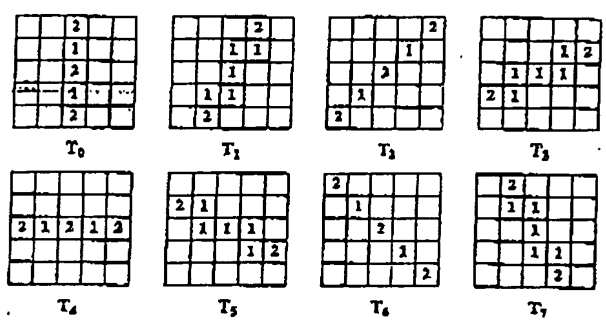
【図1】



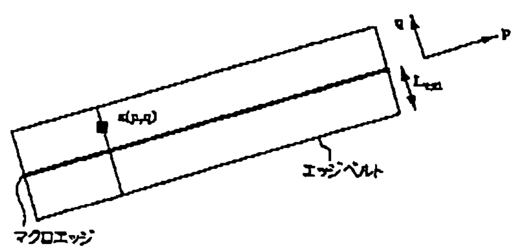

【図5】



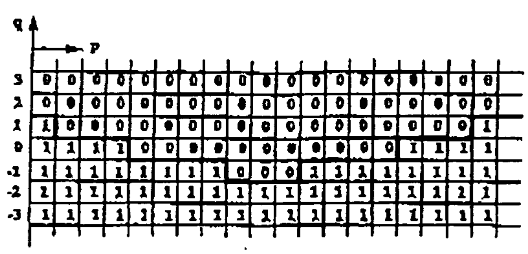
【図3】



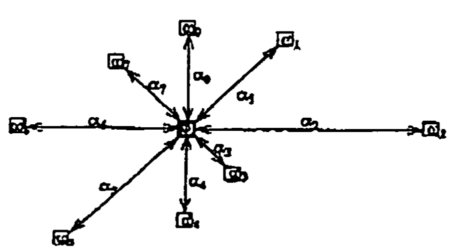

【図6】



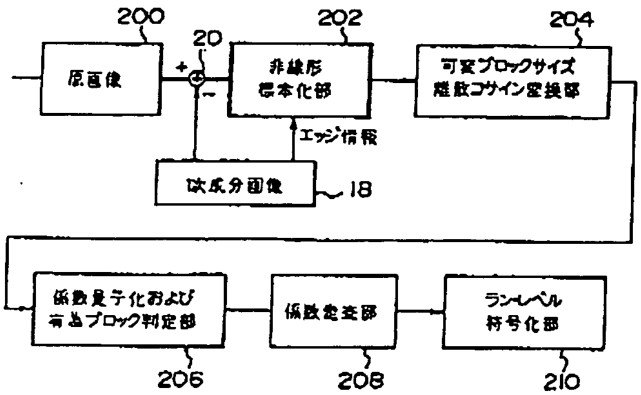

【図7】



【図8】



【図9】




可変ブロックサイズ符号化法

【図10】

【図11】

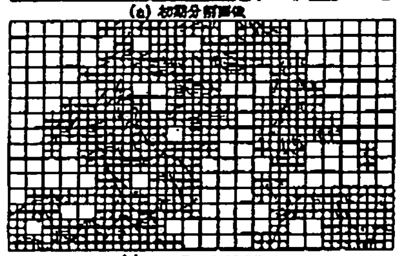
【図12】

【図13】

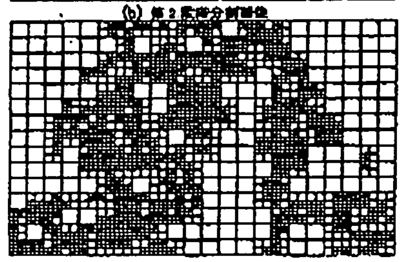
【図15】

【図14】

(a) 初期分割画像



(b) 第2段階分割画像



(c) 最終分割画像

エッジ情報を利用した非矩形標本化の例

---

フロントページの続き

| (51)Int.Cl.$^6$ | 識別記号 | 庁内整理番号 | F I | 技術表示箇所 |
|---|---|---|---|---|
| H04N 5/208 | | | | |